

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated:

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rnp) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any SIFC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:24:44 ; Search time 234 Seconds

(without alignments)
1320.604 Million cell updates/sec

Title: US-10-788-746-6

Perfect score: 2316

Sequence: 1 AVSTSPAAFYDTQTLLAAHAA.....NPPLTALTAACCPGARA 438

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 70528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	2316	100.0	1284	2	Q9WJ35 V1R1R
2	2316	100.0	1320	2	Q8J1X0 salmon panc
3	2234	96.5	1319	2	Q5JAM4 salmon panc
4	2225	96.1	1319	2	Q5JAM6 salmon panc
5	2224	96.0	1319	2	Q5JAM9 salmon panc
6	2203	95.1	1322	2	Q8Q152 sleeping di
7	2203	95.1	1324	2	Q9Q1V0 sleeping di
8	859	37.1	167	2	Q9YQ99 sleeping di
9	501	21.6	1258	2	Q8Q827 middeburg
10	499	21.5	1244	2	Q8Q828 chikungunya
11	487	21.0	1241	2	Q5WQY5 chikungunya
12	477	20.6	1241	2	Q8Q821 chikungunya
13	473	20.4	1241	2	Q8Q825 chikungunya
14	471	20.3	1248	2	Q5KXP3 chikungunya
15	470	20.3	1247	2	Q5KXP9 chikungunya
16	470	20.3	1248	2	Q5KXP9 chikungunya
17	468	20.2	1248	2	Q8JUX5 chikungunya
18	465	20.1	1253	2	Q5Y388 getah virus
19	465	20.1	1247	2	Q5Y388 getah virus
20	462.5	20.0	1253	2	Q8Q831 getah virus
21	458.5	19.8	1241	2	Q8Q847 highland j
22	456.5	19.7	1241	2	Q8Q847 highland j
23	454.5	19.6	1240	1	POL5_EBEV
24	452.5	19.5	1239	2	P89946 baemah fore
25	452.5	19.5	1239	2	P89946 baemah fore
26	452.5	19.5	1242	2	Q88795 eastern equ
27	452.5	19.5	1242	2	Q88792 eastern equ
28	452.5	19.5	1242	2	Q88790 eastern equ
29	452.5	19.5	1242	2	Q88790 eastern equ
30	452.5	19.5	1242	2	Q9PZX6 eastern equ
31	452.5	19.5	1242	2	Q88798 eastern equ

32	452.5	19.5	1242	2	Q88678_EBEV	Q88678 eastern equ
33	452.5	19.5	1242	2	Q4QXJ9_EBEV	Q4QXJ9 eastern equ
34	452.5	19.5	1242	2	Q4QXJ7_EBEV	Q4QXJ7 eastern equ
35	452	19.5	1247	1	POL5_ONNVG	P22056 O'nyong-nyo
36	450.5	19.5	1247	2	Q88794_EBEV	Q88794 eastern equ
37	450.5	19.5	1242	2	Q88793_EBEV	Q88793 eastern equ
38	449.5	19.4	1242	2	Q88796_EBEV	Q88796 eastern equ
39	449.5	19.4	1242	2	Q88797_EBEV	Q88797 eastern equ
40	447.5	19.3	1253	2	Q8Q833_V1R1R	Q8Q833 eastern equ
41	446.5	19.3	1242	2	Q9PZX7_EBEV	Q9PZX7 eastern equ
42	445.5	19.2	1242	2	Q9PZX1_EBEV	Q9PZX1 eastern equ
43	445.5	19.2	1253	2	Q9JGK8_V1R1R	Q9JGK8 eastern equ
44	442.5	19.1	1242	2	Q9PZX5_EBEV	Q9PZX5 eastern equ
45	442.5	19.1	1242	2	Q08359_EBEV	Q08359 eastern equ

ALIGNMENTS

RESULT 1	Q9WJ35 V1R1R	PRT; 1284 AA.
ID	Q9WJ35 V1R1R	PRELIMINARY;
AC	Q9WJ35	
DT	01-NOV-1999 (TREMBLrel. 12, Created)	
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)	
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)	
DE	Polyprotein.	
OS	Salmon pancreas disease virus.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;	
OC	Alphavirus.	
OX	NCBI_TaxID=84589;	
RM	[1]	
RP	NEUTROIDE SEQUENCE.	
RX	MEDLINE=99208813; PubMed=10191183; DOI=10.1006/viro.1999.9654;	
RA	Weston J.H., Welsh M.D., McLoughlin M.F., Todd D.;	
RT	"Salmon Pancreas Disease Virus, an alphavirus infecting farmed	
RT	Atlantic Salmon (Salmo salar).";	
RL	Virology 256:188-195 (1999).	
RL	EMBL; AJ012631; CAB42823.1; -; Genomic RNA.	
DR	HSSP; P03155; IYCP.	
DR	GO; GO:0016020; C:membrane; IEA.	
DR	GO; GO:0019028; C:viral capsid; IEA.	
DR	GO; GO:0008233; F:peptidase activity; IEA.	
DR	GO; GO:0004252; F:serine-type endopeptidase activity; IEA.	
DR	GO; GO:0005198; F:structural molecule activity; IEA.	
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.	
DR	InterPro; IPR002548; Alpha_E1_glycop.	
DR	InterPro; IPR000936; Alpha_E2_glycop.	
DR	InterPro; IPR000533; Alpha_E3_glycop.	
DR	InterPro; IPR000930; Peptidase_S3.	
DR	InterPro; IPR011998; Viral_glycoB_cen.	
DR	Pfam; PF00944; Alpha_core; 1.	
DR	Pfam; PF01589; Alpha_E1_glycop; 1.	
DR	Pfam; PF00944; Alpha_E2_glycop; 1.	
DR	Pfam; PF01563; Alpha_E3_glycop; 1.	
DR	PRINTS; PR00798; TOGAVIRIN.	
KW	Capsid protein; Hydrolase; Polyprotein; Protease; Structural protein.	
SEQ	SEQUENCE 1284 AA; 138915 MM; 4966A21EB019BFC0 CRC64;	
Query Match	100.0%; Score 2316; DB 2; Length 1284;	
Best Local Similarity	100.0%; Pred. No. 1.8e-171;	
Matches	438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AVSTSPAAFYDTQTLLAAHAAAPYRANCPCDGTACISPIAIDVSSGSHVLMRFGS 60	
DB	354 AVSTSPAAFYDTQTLLAAHAAAPYRANCPCDGTACISPIAIDVSSGSHVLMRFGS 413	
QY	61 OSQGTATGAGAGSTLYIRGDKVHAADNTRLVRTAKCDVQATGHTIILANCPVQGS 120	
DB	414 OSQGTATGAGAGSTLYIRGDKVHAADNTRLVRTAKCDVQATGHTIILANCPVQGS 473	
QY	121 LTVATLDGTGHOCTTVEHGVTEKFTREBSKHHLSMTWKCTRFSTTPKSAIYLVYV 180	

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:31:58 ; Search time 48 Seconds
(without alignments)
754.416 Million cell updates/sec

Title: US-10-788-746-6

Perfect score: 2316
Sequence: 1 AVSTSPAPAFDTOTIAAHAA.....NPPLTALTAACITPGARA 438

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 8265679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	438	2	US-09-674-866A-6 Sequence 6, Appli
2	2316	100.0	1359	2	US-09-674-866A-3 Sequence 3, Appli
3	437.5	18.9	1236	2	US-10-023-649A-4 Sequence 4, Appli
4	437.5	18.9	1236	2	US-10-023-649A-6 Sequence 6, Appli
5	402.5	17.4	1253	1	US-07-920-281C-3 Sequence 3, Appli
6	402.5	17.4	1253	1	US-08-466-277-3 Sequence 3, Appli
7	402.5	17.4	1253	1	US-09-688-842-3 Sequence 3, Appli
8	395.5	17.1	1245	1	US-08-801-263A-6 Sequence 6, Appli
9	395.5	17.1	1245	1	US-09-103-248-6 Sequence 6, Appli
10	395.5	17.1	1245	2	US-09-367-764-6 Sequence 6, Appli
11	394	17.0	981	2	US-09-991-258-13 Sequence 13, Appli
12	389.5	16.8	1245	1	US-08-801-263A-3 Sequence 3, Appli
13	389.5	16.8	1245	1	US-09-102-248-3 Sequence 3, Appli
14	389.5	16.8	1245	2	US-09-367-764-3 Sequence 3, Appli
15	383.5	16.6	1245	1	US-08-801-263A-10 Sequence 10, Appli
16	383.5	16.6	1245	1	US-09-102-248-10 Sequence 10, Appli
17	383.5	16.6	1245	1	US-09-367-764-10 Sequence 10, Appli
18	113	4.9	2315	2	US-09-543-681A-5434 Sequence 5434, Ap
19	107	4.6	437	2	US-09-830-189C-2 Sequence 2, Appli
20	107	4.6	437	2	US-09-991-181-355 Sequence 355, App
21	107	4.6	437	2	US-09-990-444-355 Sequence 355, App
22	107	4.6	437	2	US-10-033-301-16 Sequence 16, Appli
23	107	4.6	437	2	US-09-997-333-355 Sequence 355, App
24	107	4.6	437	2	US-09-992-598-355 Sequence 355, App
25	103.5	4.5	298	2	US-09-073-569-2 Sequence 2, Appli
26	103.5	4.5	298	2	US-09-248-796A-26762 Sequence 26762, A
27	99.5	4.3	685	2	US-10-104-047-2916 Sequence 2916, Ap

28	99.5	4.3	985	2	US-09-993-777-6 Sequence 6, Appli
29	99.5	4.3	985	2	US-09-993-777-6 Sequence 66, Appli
30	99.5	4.3	985	4	PCT-US96-03916-6 Sequence 6, Appli
31	99.5	4.3	985	4	PCT-US96-03916-6 Sequence 66, Appli
32	99.5	4.3	503	2	US-09-487-558B-424 Sequence 424, App
33	98.5	4.3	495	2	US-09-949-016-7846 Sequence 7846, App
34	97.5	4.2	1398	2	US-09-900-920-60 Sequence 60, Appli
35	97.5	4.2	1398	2	US-09-543-681A-5641 Sequence 5641, Appli
36	96.5	4.2	1402	2	US-09-125-635-12 Sequence 12, Appli
37	96	4.1	1191	2	US-09-248-796A-16243 Sequence 16243, A
38	95	4.1	382	2	US-09-603-708A-40 Sequence 40, Appli
39	95	4.1	382	2	US-09-702-705-334 Sequence 334, App
40	95	4.1	582	2	US-09-736-457-334 Sequence 334, App
41	95	4.1	582	2	US-09-614-124B-334 Sequence 334, App
42	95	4.1	582	2	US-09-671-325-334 Sequence 334, App
43	95	4.1	582	2	US-09-589-184-334 Sequence 334, App
44	95	4.1	582	2	US-09-658-824-334 Sequence 334, App
45	95	4.1	582	2	US-10-017-754-334 Sequence 334, App

ALIGNMENTS

RESULT 1					
US-09-674-866A-6					
Sequence 6, Application US/09674866A					
Patent No. 6719980					
GENERAL INFORMATION:					
APPLICANT: Akzo NO. 6719980el NV					
TITLE OF INVENTION: Structural Proteins of Fish Pancreatic Disease Virus					
FILE REFERENCE: 1/98376US					
CURRENT APPLICATION NUMBER: US/09/674,866A					
CURRENT FILING DATE: 2002-06-17					
PRIOR APPLICATION NUMBER: EP98201461.5					
PRIOR FILING DATE: 1998-05-08					
NUMBER OF SEQ ID NOS: 15					
SOFTWARE: PatentIn Ver. 2.1					
SEQ ID NO 6					
LENGTH: 438					
TYPE: PRT					
ORGANISM: Salmon pancreatic disease virus					
FEATURES:					
OTHER INFORMATION: B2					
US-09-674-866A-6					
Query Match					
Best Local Similarity 100.0%; Score 2316; DB 2; Length 438;					
Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	AVSTSPAPAFDTOTIAAHAAAPRAYCPDCTACISPAIDVSSGSDHVRMRVGS	60		
DB	1	AVSTSPAPAFDTOTIAAHAAAPRAYCPDCTACISPAIDVSSGSDHVRMRVGS	60		
QY	61	QSGVTAGGAAGETSLRYLGRDGVKHAADNTRLVRTTAKCDVLAQNGHYLANCPVQS	120		
DB	61	QSGVTAGGAAGETSLRYLGRDGVKHAADNTRLVRTTAKCDVLAQNGHYLANCPVQS	120		
QY	121	LTVAATLDTGRHOCTVFEHQVTEKTRERSKGHLSDMTKCTRFSTTPKSAALYLVDV	180		
DB	121	LTVAATLDTGRHOCTVFEHQVTEKTRERSKGHLSDMTKCTRFSTTPKSAALYLVDV	180		
QY	181	YDALPIVEISTVTCSDQCTVAVPVGTVKFKKKSADSAVTFTSDQTCBSPV	240		
DB	181	YDALPIVEISTVTCSDQCTVAVPVGTVKFKKKSADSAVTFTSDQTCBSPV	240		
QY	241	LTAAITQGPRLSAMLPGSGKEVKARIPFPPTATCRVSAVAPLPSITYESDVLAA	300		
DB	241	LTAAITQGPRLSAMLPGSGKEVKARIPFPPTATCRVSAVAPLPSITYESDVLAA	300		
QY	301	GTATCPVLLTTRNIGFHSNATSEMIQCKYLRIPVTPQGLTGNNAFMHFGSSVRYAS	360		
DB	301	GTATCPVLLTTRNIGFHSNATSEMIQCKYLRIPVTPQGLTGNNAFMHFGSSVRYAS	360		

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:32:50 / Search time 172 Seconds

(without alignments)
1064.007 Million cell updates/sec

Title: US-10-788-746-6

Perfect score: 2316
Sequence: 1 AVSTSPAAFYDTQILAAHAA.....NPPILATLALCCIPGARA 438

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1667569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1667569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	438	5	US-10-788-746-6
2	2316	100.0	1359	5	US-10-788-746-3
3	446.5	19.3	487	5	US-10-811-353-8
4	441.5	18.1	1254	5	US-10-811-353-5
5	437.5	18.9	1236	4	US-10-023-649-4
6	437.5	18.9	1236	4	US-10-023-649-6
7	437.5	18.9	1236	5	US-10-921-868A-4
8	437.5	18.9	1236	5	US-10-921-868A-6
9	402.5	17.4	1253	3	US-09-901-106-3
10	394	17.0	981	3	US-09-991-258-13
11	394	17.0	981	5	US-10-929-234-13
12	121	5.2	594	4	US-10-437-963-173742
13	113	4.9	2000	4	US-10-010-901-29
14	113	4.9	4262	5	US-10-704-781-4
15	113	4.9	4493	5	US-10-704-781-3
16	109.5	4.7	472	5	US-10-739-930-5927
17	107	4.6	437	3	US-09-989-722-355
18	107	4.6	437	3	US-09-989-723-355
19	107	4.6	437	3	US-09-989-727-355
20	107	4.6	437	3	US-09-989-727-355
21	107	4.6	437	3	US-09-989-731-355
22	107	4.6	437	3	US-09-989-732-355
23	107	4.6	437	3	US-09-991-073-355
24	107	4.6	437	3	US-09-990-442-355
25	107	4.6	437	3	US-09-991-163-355
26	107	4.6	437	3	US-09-993-604-355
27	107	4.6	437	3	US-09-990-456-355

28	107	4.6	437	3	US-09-989-721-355	Sequence 355, App
29	107	4.6	437	3	US-09-992-598-355	Sequence 355, App
30	107	4.6	437	3	US-09-989-293A-355	Sequence 355, App
31	107	4.6	437	3	US-09-989-735-355	Sequence 355, App
32	107	4.6	437	3	US-09-990-444-355	Sequence 355, App
33	107	4.6	437	3	US-09-991-181-355	Sequence 355, App
34	107	4.6	437	3	US-09-989-730-355	Sequence 355, App
35	107	4.6	437	3	US-09-990-436-355	Sequence 355, App
36	107	4.6	437	3	US-09-993-687-355	Sequence 355, App
37	107	4.6	437	3	US-09-989-734-355	Sequence 355, App
38	107	4.6	437	3	US-09-997-653-355	Sequence 355, App
39	107	4.6	437	3	US-09-989-724-355	Sequence 355, App
40	107	4.6	437	3	US-09-989-728-355	Sequence 355, App
41	107	4.6	437	3	US-09-990-441-355	Sequence 355, App
42	107	4.6	437	3	US-09-993-667-355	Sequence 355, App
43	107	4.6	437	3	US-09-997-428-355	Sequence 355, App
44	107	4.6	437	3	US-09-997-666-355	Sequence 355, App
45	107	4.6	437	3	US-09-990-438-355	Sequence 355, App

ALIGNMENTS

RESULT 1
US-10-788-746-6
Sequence 6, Application US/10788746
Publication No. US20040258707A1
GENERAL INFORMATION:
APPLICANT: Akzo Nobel NV
TITLE OF INVENTION: Structural Proteins of Fish Pancreatic Disease Virus
FILE REFERENCE: I/98376US
CURRENT APPLICATION NUMBER: US/10/788,746
CURRENT FILING DATE: 2004-02-26
PRIOR APPLICATION NUMBER: US/09/674,866
PRIOR FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: EP98201461.5
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 438
TYPE: PRT
ORGANISM: Salmon pancreatic disease virus
FEATURE:
OTHER INFORMATION: E2
US-10-788-746-6
Query Match
Best Local Similarity 100.0%; Score 2316; DB 5; Length 438;
Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 AVSTSPAAFYDTQILAAHAAAPYRACPDGACISPIAIDEVSSGSDHVLAMRVGS 60
1 AVSTSPAAFYDTQILAAHAAAPYRACPDGACISPIAIDEVSSGSDHVLAMRVGS 60
OSGTVAKGAAAGERSRLTGGDKVHADNRLVYRTAKCDVLAATGHTYLIANCPVGS 120
OSGTVAKGAAAGERSRLTGGDKVHADNRLVYRTAKCDVLAATGHTYLIANCPVGS 120
OSGTVAKGAAAGERSRLTGGDKVHADNRLVYRTAKCDVLAATGHTYLIANCPVGS 120
LTVAAATLGGTHOCTVEHGVTEKTRERKSHLSDMTKKCTFSTTPKKSLALYDV 180
LTVAAATLGGTHOCTVEHGVTEKTRERKSHLSDMTKKCTFSTTPKKSLALYDV 180
LTVAAATLGGTHOCTVEHGVTEKTRERKSHLSDMTKKCTFSTTPKKSLALYDV 180
YDALPISVETSTVTCSDQCTVRVPPGVTVKPKCKSADSAITFTSISQTFPCBEV 240
YDALPISVETSTVTCSDQCTVRVPPGVTVKPKCKSADSAITFTSISQTFPCBEV 240
YDALPISVETSTVTCSDQCTVRVPPGVTVKPKCKSADSAITFTSISQTFPCBEV 240
LTAASITGKPHLSAMTPSGGKVKARIPPPPEPTATGCVSVAPLSITTEESDVLTA 300
LTAASITGKPHLSAMTPSGGKVKARIPPPPEPTATGCVSVAPLSITTEESDVLTA 300
LTAASITGKPHLSAMTPSGGKVKARIPPPPEPTATGCVSVAPLSITTEESDVLTA 300
GTAKPVALLTRNLGFHSNATSEWIOGKYLRIIPVTPGIBLITWGNANPMHFWSSVRYAS 360

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:33:40 ; Search time 27 Seconds

(without alignments)
713.625 Million cell updates/sec

Title: US-10-788-746-6

Perfect score: 2316
Sequence: 1 AVSTSPAFYDTQILAAHAA.....NPPPLTALTAACCPGARA 438

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225428 seqs, 44002918 residues

Total number of hits satisfying chosen parameters: 225428

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA New:
1: /SIDS5/prodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /SIDS5/prodata/1/pubpaa/US06_NEW_PUB.pep.*
3: /SIDS5/prodata/1/pubpaa/US07_NEW_PUB.pep.*
4: /SIDS5/prodata/1/pubpaa/ECT_NEW_PUB.pep.*
5: /SIDS5/prodata/1/pubpaa/US09_NEW_PUB.pep.*
6: /SIDS5/prodata/1/pubpaa/US10_NEW_PUB.pep.*
7: /SIDS5/prodata/1/pubpaa/US11_NEW_PUB.pep.*
8: /SIDS5/prodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107.5	4.6	5291	US-11-052-554A-281	Sequence 281, App
2	107	4.6	437	US-10-131-826A-466	Sequence 466, App
3	107	4.6	437	US-10-973-115B-466	Sequence 466, App
4	107	4.6	437	US-10-213-535-16	Sequence 16, App
5	107	4.6	437	US-10-218-784-162	Sequence 162, App
6	107	4.6	437	US-10-219-061-162	Sequence 162, App
7	107	4.6	437	US-10-219-062-162	Sequence 162, App
8	107	4.6	437	US-10-219-064-162	Sequence 162, App
9	107	4.6	437	US-10-233-134-162	Sequence 162, App
10	107	4.6	437	US-11-290-153-466	Sequence 466, App
11	99.5	4.3	685	US-11-072-512-2916	Sequence 2916, App
12	98.5	4.3	487	US-10-745-586-11	Sequence 11, App
13	97	4.2	403	US-10-453-372-1170	Sequence 1170, App
14	96	4.1	381	US-11-087-099-3485	Sequence 3485, App
15	96	4.1	940	US-11-045-004-39	Sequence 39, App
16	95	4.1	382	US-10-703-799B-40	Sequence 40, App
17	95	4.1	1709	US-10-995-561-973	Sequence 973, App
18	95	4.1	1709	US-10-453-372-410	Sequence 410, App
19	94.5	4.1	334	US-11-096-568A-11273	Sequence 11273, App
20	94.5	4.1	365	US-11-096-568A-11272	Sequence 11272, App
21	94.5	4.1	419	US-11-096-568A-11271	Sequence 11271, App
22	94.5	4.1	583	US-11-080-991-64	Sequence 64, App
23	94	4.1	626	US-11-045-004-25	Sequence 25, App
24	94	4.1	1700	US-10-453-372-398	Sequence 398, App
25	94	4.1	1700	US-10-453-372-412	Sequence 412, App

26	94	4.1	1700	US-10-453-372-414	Sequence 414, App
27	94	4.1	1700	US-10-453-372-418	Sequence 418, App
28	93.5	4.0	859	US-11-188-743-16	Sequence 16, App
29	93.5	4.0	859	US-11-007-428-5	Sequence 5, App
30	93.5	4.0	859	US-11-183-294-14	Sequence 14, App
31	93.5	4.0	859	US-11-235-037-7	Sequence 7, App
32	93	4.0	701	US-11-188-398-13650	Sequence 13650, App
33	93	4.0	1700	US-10-453-372-416	Sequence 416, App
34	92.5	4.0	3132	US-11-087-099-1245	Sequence 1245, App
35	92	4.0	1075	US-11-100-640-12	Sequence 12, App
36	92	4.0	1709	US-11-135-855-35	Sequence 35, App
37	91.5	4.0	347	US-11-221-203-12	Sequence 12, App
38	91.5	4.0	799	US-10-493-909-8	Sequence 8, App
39	91.5	4.0	822	US-10-493-909-48	Sequence 48, App
40	91.5	4.0	1571	US-11-052-554A-2	Sequence 2, App
41	91	3.9	1694	US-11-135-855-36	Sequence 36, App
42	90	3.9	934	US-10-453-372-1158	Sequence 1158, App
43	89.5	3.9	1461	US-11-052-554A-283	Sequence 283, App
44	89	3.8	619	US-10-999-886-3	Sequence 3, App
45	89	3.8	619	US-11-136-244-3	Sequence 3, App

ALIGNMENTS

RESULT 1	US-11-052-554A-281
Sequence 281, Application US/11052554A	
Publication No. US2005028866A1	
GENERAL INFORMATION:	
APPLICANT: Sachdeva, et al.	
TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE	
FILE REFERENCE: 30853/40359A	
CURRENT APPLICATION NUMBER: US/11/052,554A	
CURRENT FILING DATE: 2005-02-07	
PRIOR FILING DATE: 2004-07-20	
PRIOR APPLICATION NUMBER: US 60/589,227	
PRIOR FILING DATE: 2004-02-06	
NUMBER OF SEQ ID NOS: 763	
SOFTWARE: Patent in version 3.3	
SEQ ID NO 281	
LENGTH: 5291	
TYPE: PRT	
ORGANISM: Escherichia coli 0157:H7	
US-11-052-554A-281	
Query Match	4.6%; Score 107.5; DB 7; Length 5291;
Best Local Similarity	19.8%; Pred. No. 1.7;
Matches 104; Conservative	68; Mismatches 175; Indels 177; Gaps 22;
QY	2 VSTSPAFYDTQILAAHAAAPRAYCP---DCBGTACISPIADEVSSGSDHVLNR 57
DB	1902 VSAPNAGSVTTAAAGSTAGNPSVTHPTVDLSAVAVSINALTADVDVNAEKAAALTL 1961
QY	58 VGSQSGTAK-----CGAAGETSLRYIG-----RDGKVA----- 87
DB	1962 SGSTSGVAGAGTCTTTTGGKTSATVANGSWSTSVPAADMAALRDGASDAQASVNVNG 2021
QY	88 -----ADNRLVVRTAKCDVLAQ-----TGHYILANCPVGSGLTVAAT 126
DB	2022 NSATTTHAVSDASAPVTITNTIAGDDITLNAABGAALITGS-----STRAAGQTVTV--T 2076
QY	127 LDGTRHQQTT-----VFEHOYKFTBRSKGHLSMTTKC 163
DB	2077 LNTGTYTGTVDGWSVSVPSADLSTLTANVTYNAVSDKAGNPASVNHLLT----- 2130
QY	164 TRFSTTPKSKALVVDVADALPISVEISTV-----VTCSDSQCTVVRPKTTPKPKCKS 219
DB	2131 -----VDTSVPV--VTINTVAGDDVYNAKHAQAQITGSSAT-----GA 2167
QY	220 ADSATVTFSDSQFT-----CREPYLTASITQGRHLASAMLPSSGKEVKARI 269

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:24:09 ; Search time 190 Seconds
(without alignments)
1012.884 Million cell updates/sec

Title: US-10-788-746-6

Perfect score: 2316
Sequence: 1 AVSTSPAPFDTQTILAHAA.....NPPPLTALTAACCPGARA 438

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 43978781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	446.5	19.3	487	ADZ75898	AdZ75898 Ross rive
2	446.5	19.3	985	AA770073	AA770073 Ross rive
3	441.5	19.1	1254	ADZ75895	AdZ75895 Ross rive
4	395.5	17.1	423	AAW70477	AAW70477 Girdwood
5	394	17.0	981	AAE18306	AAE18306 Venezuela
6	393.5	17.0	1253	AAZ28337	AAZ28337 SFV4 stru
7	389.5	16.8	423	AAW70470	AAW70470 South Afr
8	383.5	16.6	423	AAW70484	AAW70484 Sindbis v
9	127.5	5.5	456	ADCO1828	ADCO1828 C. albica
10	113	4.9	2000	ADP05149	ADP05149 Bacterial
11	113	4.9	2315	ADP05149	ADP05149 Bacterial
12	113	4.9	4262	AAO08149	AAO08149 Human muc
13	113	4.9	4493	AAO08148	AAO08148 Human muc
14	110	4.7	437	AA787261	AA787261 Human big
15	110	4.7	998	ABU53140	ABU53140 Human tes
16	109.5	4.7	472	ADT55850	ADT55850 Plant pol
17	108	4.7	437	AA680767	AA680767 Human gra
18	107.5	4.6	5291	ADCO1014	ADCO1014 Enterohae
19	107.5	4.6	5291	AAE91571	AAE91571 Microbial
20	107	4.6	437	AA787750	AA787750 Human PRV
21	107	4.6	437	AA787750	AA787750 Membrane
22	107	4.6	437	AA787750	AA787750 Membrane
23	107	4.6	437	AA787750	AA787750 Membrane
24	107	4.6	437	AA787750	AA787750 Membrane

25	107	4.6	437	4	AA655260	AA655260 Human PRO
26	107	4.6	437	5	AAU83672	AAU83672 Human PRO
27	107	4.6	437	5	ABG92707	ABG92707 Human sec
28	107	4.6	437	5	ABG91359	ABG91359 Novel hum
29	107	4.6	437	5	ABG31399	ABG31399 Human PRO
30	107	4.6	437	5	ADY31902	ADY31902 Novel hum
31	107	4.6	437	6	ABU72375	ABU72375 Novel hum
32	107	4.6	437	6	ABU58075	ABU58075 Human PRO
33	107	4.6	437	6	ABU59153	ABU59153 Novel hum
34	107	4.6	437	6	ABU80865	ABU80865 Human sec
35	107	4.6	437	6	ABU82665	ABU82665 Human sec
36	107	4.6	437	6	ABU17848	ABU17848 Novel hum
37	107	4.6	437	6	ABU60584	ABU60584 Human sec
38	107	4.6	437	6	ABU80819	ABU80819 Human PRO
39	107	4.6	437	6	ABO33785	ABO33785 Novel hum
40	107	4.6	437	6	ABU13966	ABU13966 Human PRO
41	107	4.6	437	6	ABU81102	ABU81102 Human PRO
42	107	4.6	437	6	ABU72551	ABU72551 Novel hum
43	107	4.6	437	6	ABU66802	ABU66802 Human PRO
44	107	4.6	437	6	ABG73312	ABG73312 Human PRO
45	107	4.6	437	6	ABU59883	ABU59883 Novel sec

ALIGNMENTS

RESULT 1	ADZ75898	ADZ75898 standard; protein; 487 AA.
ID	ADZ75898	ADZ75898 standard; protein; 487 AA.
AC	ADZ75898;	
DT	28-JUN-2005	(first entry)
DE	Ross river virus E3-E2 protein, seq id 8.	
DE	Ross river virus E3-E2 protein, seq id 8.	
KW	Gene therapy; retrovirus based gene therapy; cystic fibrosis; airway epithelial cell; E3-E2.	
OS	Ross River virus.	
PN	US2005112098-A1.	
XX	26-MAY-2005.	
PD	26-MAY-2005.	
PF	26-MAR-2004; 2004US-00811353.	
XX	26-OCT-2001; 2001US-0353221P.	
PR	26-OCT-2001; 2001US-0353221P.	
PR	04-JUN-2002; 2002US-0356436P.	
PR	28-OCT-2002; 2002US-0356436P.	
PR	27-MAR-2003; 2003US-0458070P.	
PR	04-JUN-2003; 2003US-0458070P.	
XX	(PUSD) PURDUE RES FOUND.	
PA	(IOWA) UNIV IOWA RES FOUND.	
PI	McCray PB, Sanders DA, Jeffers SA, Davidson BL, Sinn PL; WPI; 2005-384668/39.	
DR	N-PSDB; ADZ75898.	
XX	Introducing a selected nucleic acid sequence into an airway epithelial cell by contacting the airway epithelial cell with a pseudotyped retrovirus comprising a glycoprotein and retroviral capsid comprising the nucleic acid sequence.	
PT	Introducing a selected nucleic acid sequence into an airway epithelial cell by contacting the airway epithelial cell with a pseudotyped retrovirus comprising a glycoprotein and retroviral capsid comprising the nucleic acid sequence.	
XX	Disclosure; SEQ ID NO 8; 102pp; English.	
PS	The invention relates to introducing a selected nucleic acid sequence into an airway epithelial cell. The method comprises contacting the airway epithelial cell with a pseudotyped retrovirus, where the pseudotyped retrovirus comprises a glycoprotein in which a portion of an	

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:28:04 ; Search time 42 Seconds
(without alignments)
1003.403 Million cell updates/sec

Title: US-10-788-746-6

Perfect score: 2316

Sequence: 1 AVGSTPAAPFDQTGLAAHAA.....NPPPLVLTALACCTPGARA 438

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR.80:*
2: PIR1:*
3: PIR2:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	456.5	19.7	1241	2 S26373	genome polyprotein
2	454.5	19.6	1240	1 VHWVWV	structural polypro
3	452.5	19.5	1242	2 S72350	structural polypro
4	452	19.5	1247	1 VHWVW2	structural polypro
5	446.5	19.3	422	1 VHWVW48	structural polypro
6	442.5	19.1	1242	2 A56605	structural polypro
7	437.5	18.9	1239	1 VHWVWE	structural polypro
8	434.5	18.8	1236	1 VHWVWE	structural polypro
9	416.5	18.0	1254	1 VHWVRA	structural polypro
10	413.5	17.9	422	1 VHWVW70	structural polypro
11	399	17.2	1254	1 JQ1978	structural polypro
12	398.5	17.2	1253	1 VHWVW	structural polypro
13	397.5	17.2	1245	1 VHWVW82	structural polypro
14	396	17.1	1254	1 JQ1979	structural polypro
15	396	17.1	1255	1 D44213	structural polypro
16	395	17.1	1254	1 VHWVW7	structural polypro
17	394.5	17.0	755	2 S42462	structural polypro
18	393.5	17.0	1145	2 S37136	structural polypro
19	393	17.0	1254	1 VHWVW	structural polypro
20	386.5	16.7	1245	1 VHWVW	structural polypro
21	386.5	16.6	1245	1 VHWVW2	structural polypro
22	384	16.6	1255	1 B44213	structural polypro
23	383.5	16.6	423	1 VHWVW8	structural polypro
24	126	5.4	69	2 G37264	E2 glycoprotein -
25	115	5.0	69	2 F37264	E2 glycoprotein -
26	109.5	4.7	472	2 T05852	hypothetical prote
27	109	4.7	1777	2 T34369	hypothetical prote
28	107.5	4.6	5188	2 B85547	probable RTX faml
29	107.5	4.6	5291	2 F90696	hypothetical prote

30	106.5	4.6	1251	2 T21389	hypothetical prote
31	104.5	4.5	1609	2 S25345	probable membrane
32	103.5	4.5	551	2 G84301	hypothetical prote
33	103	4.4	1694	2 S50065	staiadhesin - mou
34	102.5	4.4	967	2 S66852	hypothetical prote
35	101.5	4.4	1522	2 H88380	protein T2287.3 li
36	101	4.4	587	2 JH0464	DM-GRASP precursor
37	101	4.4	588	2 JH0506	adhesion molecule
38	99	4.3	503	2 S63257	probable membrane
39	99	4.3	630	2 A39344	tumor-associated m
40	98.5	4.3	487	2 JCT126	testis zinc finger
41	98	4.2	588	2 A45254	surface glycoprote
42	97	4.2	66	2 A37264	E2 glycoprotein -
43	97	4.2	1676	2 A56508	anucleate primary
44	96.5	4.2	747	2 S18685	Sc/Sw1 protein -
45	96.5	4.2	748	2 S18690	Sc/SvQ protein - E

ALIGNMENTS

RESULT 1

S26373 genome polyprotein - eastern equine encephalomyelitis virus
N:Contains: 6X protein; capsid protein C; envelope protein E1; envelope protein E2; en
C:Species: eastern equine encephalomyelitis virus
C:Date: 06-Jan-1994 #sequence_revision 01-Nov-1996 #text_change 31-Dec-2004
C:Accession: S26373
R:Volchikov, V.E.; Volchikova, V.A.; Netesov, S.V.
Mol. Gen. Microbiol. Virusol. 5, 8-15, 1991
A:Title: Complete nucleotide sequence of the eastern equine encephalomyelitis virus ge
A:Reference number: S26369; MUID:91375524; PMID:1896061
A:Accession: S26373
A:Molecule type: mRNA
A:Residues: 1-1241 <VOL>
A:Cross-references: UNIPROT:Q66579; UNIPARC:UP100000P1639; EMBL:X63135; NID:959185; PI
A:Note: sequence could not be checked because of bad print in paper
C:Keywords: capsid protein, envelope protein, glycoprotein, polyprotein
F:1-260/Product: capsid protein C #status predicted <CNP>
F:261-323/Product: envelope protein E1 #status predicted <EP1>
F:324-743/Product: envelope protein E2 #status predicted <EP2>
F:744-800/Product: 6X protein #status predicted <6XP>
F:801-1241/Product: envelope protein E1 #status predicted <EP1>

Query Match	19.7%	Score	456.5	DB	2	Length	1241
Best local similarity	30.2%	Pred. No.	5.3e-27				
Matches	133	Conservative	64	Mismatches	189	Indels	55
							Gaps 17
QY	21	ASPRAYCPDGDGACTSPIAIDRVSSGSDHVRMRVSGSGVYAKGAGGTSIRYL-	79				
DB	335	ARPIAIDCPNGHSGHRCSPFAIBSRGDAGAVIRIQTSAPFGAKTDD-VDLAWS	389				
QY	80	---GRDGKVAHADTRLVVETTRAKDVLQATGHTLANCPVGSLLVVAATLDGRHOCTT	136				
DB	390	FMNKRTQKSLIKDN--LHVRTSAPCSLVHHGYIILAQCPGDDVTVTFHDPNRHTCTV	447				
QY	137	VFEHQV-----TEKTFERSKGNHLSDMTKCTRFSTTPPKSALLVDVYDALPISVEI	190				
DB	448	A--HKVEFRPVGRKRYRHPPEHGVEL-----PCNRY--THKRAQGHVVEHQ--FGLVAD	497				
QY	191	STVTCGSDQCVAVVPPGTVPKPKCKKSADSAVITPSSDQCTFCHEFVLTAAISITQK	250				
DB	498	HSLSLHSAAKKTIVPSSGAQVKYCKCPDVBEGT--TSSDYTTICDVKOCRAVLINRK	554				
QY	251	PHL-PSAMLPSSG--GKEVYARIPFPPEPETAICRVSVAPLPSITYESDVLTLGAKYFVL	308				
DB	555	KRVVNSGRLPKRGEDTRKGLHVFVYVAKCINTADPEPLVHKHRTLLHLVDPDHTL	614				
QY	309	LTFTNIGFHSNATSEWIOGKTLRRIPTYPGIEILTWGNANAPMFWSSVRYVSGADAYFW	368				
DB	615	LTFTSLASGDANPTQWIERPTTVNFTVGTGELBYTWGNHPPKRWMAQ--ESGEGDPHGWPH	673				